

# Automating RTI: Automatic light direction detection and correcting non-uniform lighting for more accurate surface normals

Matthew McGuigan\*, Jacqueline Christmas

*University of Exeter, Exeter, EX4 4QF, UK*

---

## Abstract

Reflectance Transformation Imaging (RTI) (Malzbender et al., 2001) is a photometric stereo technique that enables the interactive relighting of the object of interest from novel lighting directions, and an estimation of surface topography through the calculation of surface normal vectors. We propose a novel, fully automated technique for correcting common lighting errors in RTI and markedly improve the accuracy of surface normal estimation, as well as increasing the legibility of low relief surface variations. This moves RTI from the qualitative domain (e.g. enabling the reading of weathered inscriptions) into the quantitative domain of computer vision. RTI assumes only light direction, and not received intensity, changes as the object is imaged. Like other authors we show that this assumption is false and propose a novel method to correct for it. However, we estimate the lighting directions automatically, unlike other proposed correction techniques. Our method also requires no calibration equipment, meaning it can be easily retrofitted to any existing stack of RTI photographs. We increase the simplicity of the standard highlight RTI method by automatically detecting lighting directions and maintain its appeal to non-imaging professionals.

**Keywords:** Reflectance Transformation Imaging, Photometric Stereo, Image enhancement

---

## 1. Introduction

Reflectance Transformation Imaging (RTI) (Malzbender et al., 2001) is a photometric stereo technique that enables the interactive relighting of the object of interest from novel lighting directions, and an estimation of surface topography through the calculation of surface normal vectors. The method combines a stack of digital images, each of which is lit from a different direction, into a new representation (called a Polynomial Texture Map, or PTM, by (Malzbender et al., 2001)) that models how the reflectance of each pixel varies by lighting direction. RTI is widely used in the cultural heritage sector as it is relatively easy and inexpensive to perform (Mudge et al.,

2006), is supported by free software from Cultural Heritage Imaging<sup>1</sup>, and, for example, enables the viewer to reveal markings that are not legible, or even visible, from a single photograph (Boute et al., 2018).

Key to the method is knowing the lighting direction for a given photograph. If the illumination is provided by a fixed lighting dome (see Fig. 1), then these directions are defined by the hardware. However, much RTI is performed without a dome and with a hand-held flash instead, known as highlight RTI (H-RTI) which was introduced by Mudge et al. (2006). The light direction is calculated from the reflection of the flash on a shiny, black, spherical marker, or other calibration devices (Huang et al., 2015; Giachetti et al., 2018), included in the

---

\*Corresponding author: Tel.: +44-784-635-0846;

Email address: M.McGuigan@exeter.ac.uk (Matthew McGuigan)

<sup>1</sup>[http://culturalheritageimaging.org/What\\_We\\_Offer/Downloads/](http://culturalheritageimaging.org/What_We_Offer/Downloads/)

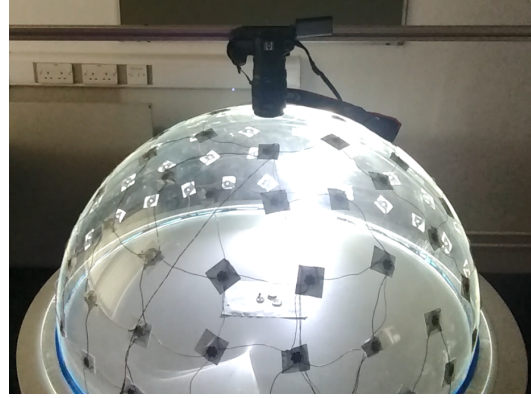
photographs. The process for creating the PTM depends on accurate light direction estimation and on the assumption that the only difference in the level of illumination in each photograph is caused by the variation in light direction.

In this paper we introduce a fully automated RTI technique that improves the accuracy of the surface normal estimation, moving RTI from the qualitative domain towards the quantitative domain. The technique requires no calibration equipment and can be retrofitted to any existing RTI image stack. The proposed method compensates for brighter and darker regions caused by relative distance to the light source, significantly reducing non-uniform lighting errors, automatically detects the light direction from each image photometrically, removing the errors introduced by the bias of the shiny spheres in high-light RTI, as well as the obstructive shadows they cause, and removes the need for reflective spheres or other calibration devices through automation, thereby increasing simplicity of RTI and its appeal to non RTI specialists.

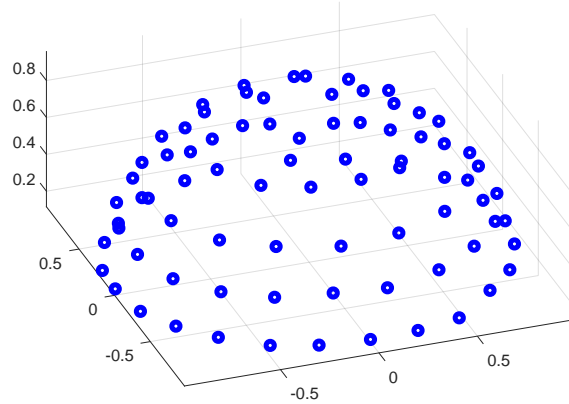
In section 2 we outline the relevant background theory behind RTI, and in section 3 we describe some of the known inaccuracies introduced by the RTI process. We present our new method for addressing these issues in section 4 and then compare the effectiveness of the new method with the standard highlight RTI method (Mudge et al. (2006)) in section 5. RTI assumes a diffuse surface and is usually intended to work for surface relief rather than significantly varying surface heights, however we will also demonstrate in section 5 that our method produces improved results on a semi-cylindrical surface. Conclusions are drawn in section 6.

## 2. Background

In RTI, the object of interest is photographed under as wide a range of lighting directions as possible. The technique assumes that the camera, viewpoint and object are in precisely the same position for each image in the stack, and that the only thing that changes in each image is the *direction* of the illumination (Malzbender et al., 2001). The interactive relighting enabled by RTI makes it possible to view the imaged object under “virtual” illumination directions that were never actually captured in the original stack of photographs. The light intensity,  $L$ , of a given



(a) RTI lighting dome.



(b) Resulting normalised light directions of the dome.

Figure 1: One type of RTI lighting set-up is (a) a hemispherical lighting dome that provides fixed light directions, with the camera at the apex. The plot in (b) shows the normalised light directions calculated for this dome.

pixel in these virtual images is constructed from a simple bi-quadratic function of the selected lighting direction (specified by the projection  $(l_u, l_v)$  of the normalised light vector onto the plane of the photograph), where the coefficients of the function  $(a_0, \dots, a_5)$  are learned from the image stack:

$$L = a_0 l_u^2 + a_1 l_v^2 + a_2 l_u l_v + a_3 l_u + a_4 l_v + a_5 \quad (1)$$

Each pixel’s luminance function is learned independently. For a given pixel, maximising  $L$  with respect to  $l_u$  and  $l_v$  enables us to calculate the surface normal vector at that

point, giving information about the shape of the surface of the object (Malzbender et al., 2001). In this section we will explain how the photographs are taken, how the lighting direction vectors are derived, the details of the standard highlight RTI technique and, briefly, some alternative methods for revealing surface topography.

### 2.1. RTI image stack acquisition

In order to obtain the RTI image stack, a camera is fixed in position (the scene must also be static) and each image is lit from a different direction (see Fig. 1), meaning that a digital image contains the reflectance response of the scene from that particular lighting direction. RTI's traditional estimation of surface normals was originally a by-product of the virtual relighting, mainly being used as a method of contrast enhancement or to make sure that the image stack was correctly constructed by inspecting the normals generated to make sure they were broadly sensible. RTI builds on the Bidirectional Reflectance Distribution Function (BRDF) defined by Edwin (1977), which estimates the intensity reflected from a material from a given viewing point and given lighting direction. In a similar manner, RTI approximates the intensity reflected from a material lit from any direction, but keeps the viewing point constant (i.e. the camera is kept stationary) which removes the need for an enormous volume of photographs to sample the viewing position space.

### 2.2. Dome based RTI

As originally proposed by Malzbender et al. (2001), data acquisition time is short when lighting for each photograph can be performed using a rigid hemispherical structure that contain tens of lights (typically between 20 and 100 lights) which fire automatically and trigger the camera; see Fig. 1. In this set-up the camera is rigidly fixed at the dome apex to capture images synchronously with the lights. The illumination directions for the RTI image stack can be predetermined before imaging since the dome is of known lighting geometry, saving the need for computing the light direction in the processing stage. These devices are known as RTI domes and can provide a data acquisition time of the order of a few seconds, however this equipment can be expensive and fragile, making it difficult to store and transport.

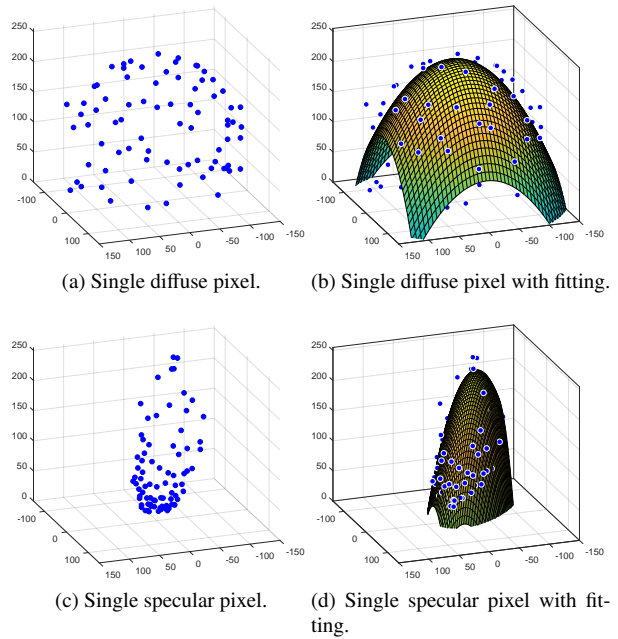


Figure 2: In RTI, a bidirectional reflectance function is fitted to the lighting intensity measurements independently for each pixel. In each of the four plots, the points represent the measurements, with the distance from the origin representing the intensity measured when the pixel is lit from that direction. The surfaces are bi-quadratic polynomial fits to those points. (a) and (b) show results for a diffuse (matt) material; the intensity distribution is approximately uniform over the hemisphere. (c) and (d) shows results obtained from a more specular material; the intensity distribution is clearly biased towards a particular direction.

### 2.3. Highlight RTI with reference spheres

The issue of portability and expense is addressed by Mudge et al. (2006), who propose a technique whereby the camera is mounted on a tripod and the light source is moved around by hand. A shiny black reflective sphere is placed in the field of view of the camera and the lighting direction is then determined in the processing stage by tracking the specular highlight caused by the light source on this sphere. Using the sphere, the lighting direction can thus be calculated intrinsically from the RTI image stack itself. This alternative method is known as Highlight RTI (H-RTI), and has aided in RTI's broad appeal to the cultural heritage community since it is inexpensive and requires less knowledge of lighting geometry.

From the image stack two crucial things are obtained:

the direction from which each image was lit,  $(l_u, l_v)$ , (Malzbender et al., 2001) and the associated light intensity emitted at each pixel,  $L$ . Note the subscripts  $u$  and  $v$  here because the light vector is normalised and projected onto the local texture coordinate system  $(u, v)$ . A reflectance distribution is a 3D measure of the spread of light, which we can now plot for any pixel, as shown in Fig. 2. In these measurement distributions the  $L$  value for a given pixel on the imaging sensor is taken from the images and is used to scale the direction vector from which the image was lit.

#### 2.4. Polynomial Texture Maps (PTMs)

We will now introduce polynomial texture maps (Malzbender et al., 2001), but to avoid confusion RTI will be referred to as the *process* from which PTMs are *output*. A PTM is a re-lightable texture map that can generate surface normal data using the coefficients from (1). RTI assumes that the surface material being imaged is diffuse and non-specular (non-shiny), we will explain what the main differences between these material types are.

Since the camera does not move during RTI, a given pixel on the image sensor will correspond to the same region on the object in each of the images. The PTM is fitted to the reflectance distribution independently for each spatial pixel in the image. In Fig. 2, we show the results of this fitting for two different materials: (a) shows the reflectance distribution measured for a diffuse (matt) material which appears spherical and homogeneous because by definition the material reflects roughly the same amount from all directions; (b) shows the same measurements superimposed on the fitted bi-quadratic surface; (c) and (d) show the same for a less diffuse (i.e. more specular or shiny) material where the distribution appears more elongated and drawn-out in one direction. Very specular surfaces reflect light only in one direction, making the distribution more like either a delta function (if that reflection direction is captured in the photographs) or a zero function (if it is not captured). Both of these cases lead to a very poor fit of the bi-quadratic approximation, which is the reason why RTI is not appropriate for very specular materials. Once the six coefficients ( $a_0$  to  $a_5$ ) for each pixel have been obtained the process can easily be reversed, and input any arbitrary light direction  $(l_u, l_v)$  to obtain the luminosity  $L$  for any point on the upper hemisphere of the unit sphere. We treat each  $(R, G, B)$  colour

channel from a source image as three separate images, and compute a different value of  $L$  for each colour channel separately, then recombine them as  $(L_R, L_G, L_B)$  to create new  $LRGB$  virtually relit images. From these coefficients, estimates of the orientation of each pixel within the image are made through computing surface normal data by differentiation in order to find the maximum.

RTI has been accessible to the cultural heritage community thanks to the availability of the *PTM Fitter* (Hewlett-Packard, 2001) together with a user interface, *RTI Builder* (Cultural Heritage Imaging, 2016), which are used to preprocess the RTI image stack to calculate the lighting directions. The *RTI Builder* is also accompanied by *RTI Viewer* (Cultural Heritage Imaging, 2016) which allows for virtual relighting and surface normal generation in an easy-to-use interface for end-users.

#### 2.5. Other reflectance functions

As previously stated in section 2, RTI fits a function on a per-pixel basis to the measured reflectance distribution at each point, which varies for each pixel depending on the orientation and material at each individual pixel. The bi-quadratic is the original and primary function still used in RTI. There exists mathematical functions which can be fitted to the reflectance distribution, each of which offer varying numbers of coefficients and fidelity. The most notable of these alternative functions are hemispherical harmonics (HSH) (Gautron et al., 2004) and discrete modal decomposition (Pitard et al., 2015). These functions offer more degrees of freedom since they have more coefficients than the standard six coefficient bi-quadratic polynomial, meaning they could in theory better fit individual pixels which are more complex (such as specular pixels). However, these functions can suffer from over-fitting and the best results are found when the fitting method is selected on a case-by-case basis, so we opt to use the bi-quadratic polynomial in our method since we will address improvements to RTI in general and is not material specific. It should be noted that the selection of a function for the per-pixel fitting (e.g. PTM, HSH etc) is independent of any light direction estimation or intensity corrections, and that an improvement in light direction estimates would improve any per-pixel fitting.

Laser scanning is a comparable, yet considerably more expensive and complex, topographical imaging technique.



It has already been shown that RTI can provide better results than laser scanning for surfaces with very low surface relief (Hammer et al., 2002).

### 3. RTI errors and related work addressing them

The standard highlight RTI method makes the inaccurate assumption that the scenes are lit by collimated light rays. We will outline these inaccuracies and some related work which has been undertaken to combat this.

#### 3.1. Known errors with highlight RTI

The inaccuracies in standard highlight RTI become more noticeable for larger artefacts and image scenes, since light rays diverge more with respect to distance. Due to the very nature of light propagation, it is impossible to artificially produce exactly collimated light (light consisting of exactly parallel rays). Some light sources are considered to be approximately collimated, when the radius of curvature of the spherical wavefront is sufficiently large (i.e. the source is sufficiently far away) that when the wavefront is incident it can be considered flat and an approximate plane wave. Laser light is more parallel than traditional light sources, but still suffers from beam divergence due to diffraction and is also impractical for RTI since a laser beam width is so small.

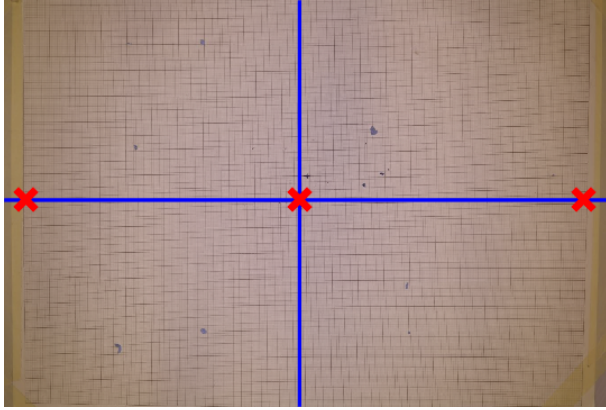
The effect of this variation in intensity in RTI is demonstrated in Fig. 3 where fitted lighting distributions are shown measured at 3 different pixels which should be identical as they have identical orientation. Their RTI polynomial fittings have been plotted. These three pixels should have the same reflectance distribution (Fig. 3c is correct for a diffuse pixel facing the camera), but it can be seen that the left and right hand side pixels (Figs. 3b and 3d) show significantly different distributions. The differences visible in the plots in Figs. 3b, 3c and 3d are in fact due to the fact they are nearer in proximity to the light source when low elevation light images are taken (and hence experience much stronger intensity). RTI somewhat naively fits a polynomial to the reflectance distribution which assumes that each pixel has been lit equally and uniformly.

This is the crux of the problem: regular light sources simply do not spread the light evenly enough. Standard highlight RTI neglects both non-uniform intensity and

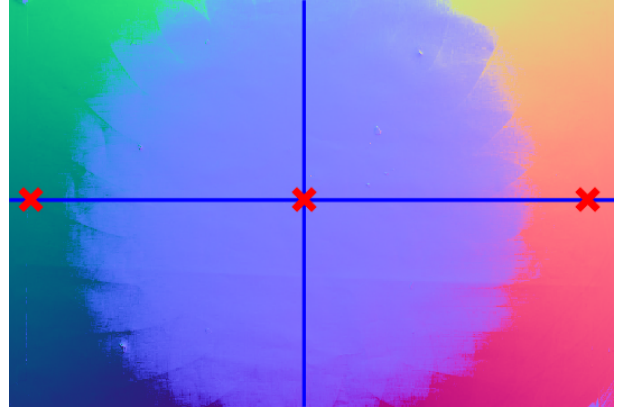
non parallel rays, which can result in extreme errors in surface normal estimates. Fig. 4a shows the surface normal data extracted from the PTM. This colour scheme is a standard form of displaying surface normal data in image form, where the  $(x, y, z)$  components of the surface normals are represented as the  $R, G, B$  colour channels respectively.

The uniformly flat and diffuse surface shown in Fig. 3 means that each pixel should reflect roughly the same amount of light and have approximately the same surface normal. We would expect that the surface normals should be uniform across the image, with a  $z$  component of 1 (as is the case in the centre of Fig. 4a), and that a few irregularities should be visible due to small pieces of material or dust present on the surface. Instead it is evident in Fig. 4a that there is a ring-like effect around the periphery of the image. This is due to non-uniform lighting, which has resulted in the pixels near the periphery of the image receiving too much illumination from the light sources they are nearest to. The RTI computations do not take this effect into consideration, and so are affected by this error caused by proximity to the light source. These images of surface normal vectors serve as useful tools in quantifying the inaccuracies caused by non-uniform illumination. A histogram of the  $z$  components of surface normals is shown in Fig. 4b, which shows a high count of  $z = 1$  surface normals but also a significantly high count of  $z = 0$  surface normals which are from the periphery of the image.

In standard highlight RTI a shiny reference sphere is placed in the scene which (to avoid obstructing the object being imaged) is never placed in the centre of the image and often placed at the periphery. This leads to the light directions being heavily biased to wherever the reference sphere is arbitrarily placed. This problem is not addressed in standard highlight RTI leading to the majority of RTI users unintentionally and unknowingly inducing errors in their light direction estimates. The sphere itself is obstructive and can cause errors in the fitting process when it casts shadows on the surface being imaged. This type of obstruction-induced error is shown in the surface normal image in Fig. 5, where a sphere is placed next to the object for RTI and the shadows have clearly degraded the surface normal estimates.



(a) Image of flat diffuse surface where the origin is at the centre.



(a) The standard surface normal image generated from Fig. 3a.

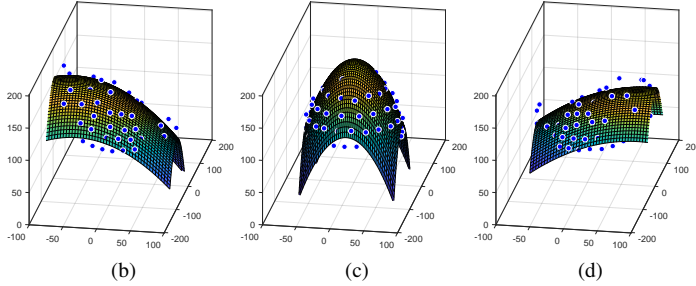


Figure 3: A demonstration of the lighting problem with RTI. (a) shows the colour image of a flat, diffuse surface (white graph paper) photographed using an RTI dome, overlaid with  $x$  and  $y$  axes. The camera is located above the origin (the centre cross); the three crosses mark the positions of three selected pixels. (c)–(d) show the reflectance distributions measured and fitted at the three selected pixel locations (left, centre and right respectively) marked by crosses in (a). For each pixel, the material and orientation are the same, but the distributions appear to be facing different directions due to different amounts of non parallel light received.

### 3.2. Related work

Given a homogeneous and diffuse surface such as white paper, this ideal case can be exploited because white paper is (approximately) Lambertian. For the case when the image subject is not conveniently Lambertian, methods have been proposed by Giachetti et al. (2018) in which 3D printed Lambertian surfaces are physically placed around the border of the scene in order to reveal the spread of illumination and then to measure this photometrically. This method successfully compensates for non-uniform illumination, but requires the insertion of extra physical appara-

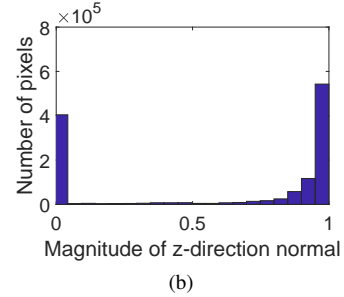


Figure 4: (a) shows the inaccurate surface normal data generated from the flat surface shown in Fig. 3a (see online for colour). One may be tempted to assume these errors are caused by vignetting, but we know this is not true because if remove low elevation light images that the effect disappears. This means the effect is in fact caused by low elevation lights in RTI which are disproportionately illuminating the borders of the image. This image should be an even blue colour; the central ring effect is caused by errors from non-uniform lighting. A histogram of the  $z$  component of the surface normals across the whole image is shown in (b). This should only show a sharp peak at  $z = 1$ , but instead we also see a peak at  $z = 0$ .

tus and steps. Furthermore, this method may be undesirable because RTI is often performed on fragile artefacts where placing such a structure in the image frame may pose additional risk of damage to an artefact, as well as the fact that this would mean more equipment must be transported.

Huang et al. (2015) introduce a promising distance-compensated pixel intensity framework that aims to correct for non-uniform lighting and estimate light directions. However, their method requires a colour-checker

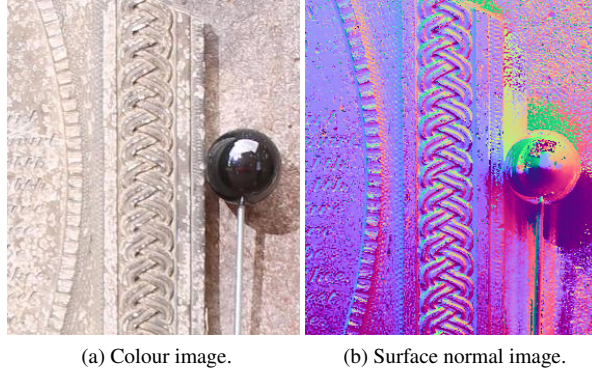


Figure 5: Highlight RTI requires a shiny reference sphere to be inserted in the scene. (a) shows the colour image of an object of interest, with the reference sphere. (b) shows the effect that shadowing from this sphere has on the calculated surface normals (for colour see online version).

calibration target to be placed in the scene, the estimation of 3D scene points, the requirement of initial values for combined albedo and vignetting, and initial surface normal estimates. This could be problematic for imperfect highlight RTI image stacks, since the method computes initial estimates of surface normals (already known to be potentially inaccurate), and obtains 3D scene point estimates by inputting them into a 3D surface reconstruction technique (Agrawal et al., 2006) which relies on sufficiently accurate surface normals to create a gradient map. This 3D surface reconstruction technique has shown very promising results with synthetic data but will only partially recover features in real images (Agrawal et al., 2006), missing more intricate details. The method would also likely suffer inaccuracies in arbitrary highlight RTI conditions such as the non ideal image stacks used in this paper where initial surface normal estimates are found to be highly inaccurate (see Fig. 17, left column). Much like Giachetti et al. (2018), the technique proposed by Huang et al. (2015) is very effective when RTI image stacks meet certain idealised criteria, but it would be impossible to retrofit the method to pre-existing RTI image stacks since these techniques require specific items to be placed in the scene. These extra steps pose potential complications and may be unattainable since the method requires the acquisition of a 3D printed structure. We propose a method that does not require such a structure, but it still relies on RTI

---

**Algorithm 1** Correcting RTI for non-uniformly lit images and improving light direction estimation using the proposed method. The input dataset is a regular RTI image stack and the result is a more accurate estimate of surface normal after corrections for non-uniformly lit images and automated light direction estimation.

---

```

for each image,  $I$ , in the image stack do
  - Generate mean-corrected version of the image,  $I_c$ .
  - Fit a bi-quadratic function to the lighting intensity across  $I_c$ .
  - Use the fitted function to compensate for over-lit and under-lit regions.
  - Automatically detect lighting direction from artificial mean-corrected image rather than using shiny reference sphere.
end for

```

---

illumination being adequately and consistently pointed by hand toward the object, as should be the case anyway.

A similar method to that introduced by Huang et al. (2015) is proposed by Winnemöller et al. (2005) who use dimensionality reduction to produce estimates of the vectorised light directions. This method relies on the distance between consecutive lighting directions in the image stack being small and requires the use of a diffuse reflector in addition to the light source. Whilst producing approximate light directions in a fast acquisition time, the required manner of lighting may further obfuscate the RTI process to non imaging science users, if, for example the distance between consecutive lighting directions is not small enough. This constraint on RTI would also likely prevent retro-fitting the technique, since highlight RTI image stacks are often captured from arbitrarily consecutive lighting directions.

For these reasons it became clear that any potential solution must be as automated as possible, with emphasis on keeping the data acquisition process as simple as possible. Hence, a novel method is proposed which corrects for non-uniform illumination, automatically detects the lighting direction and does not require any further steps for the user. The technique in fact aims to minimise the equipment needed since the method proposes a way to automatically detect the lighting direction - removing the need for shiny reference spheres and allowing retro-fitting of our technique to existing highlight RTI image stacks.

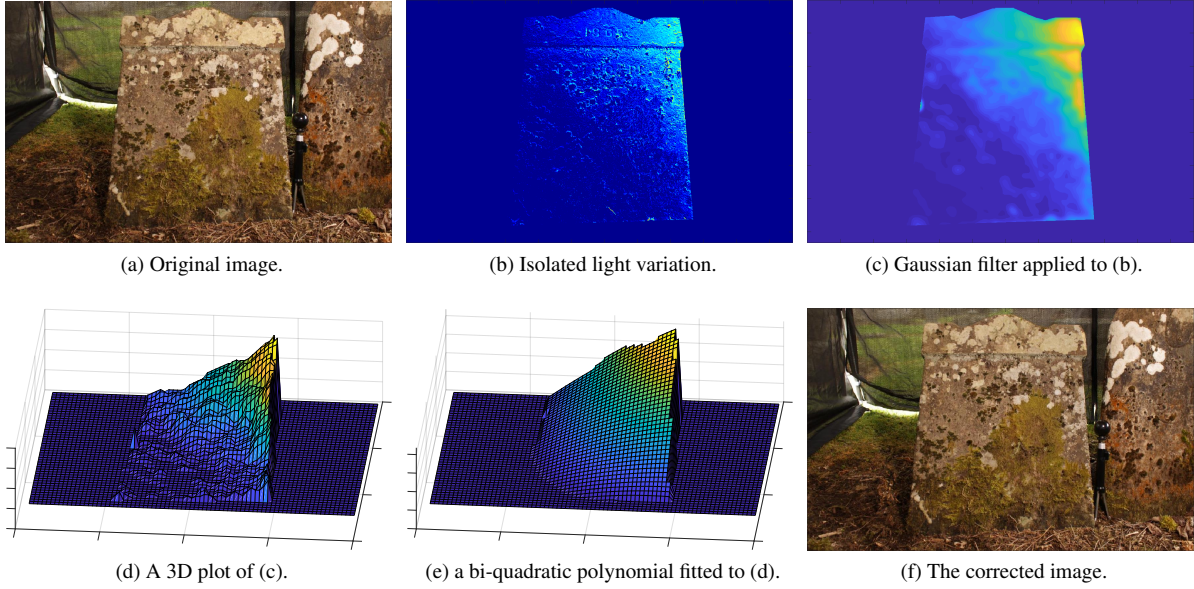


Figure 6: An outline of the proposed method applied to an RTI image stack that has been lit somewhat problematically, showing (a) the original non-uniformly lit image, (b) the light spread isolated as the image is mean-corrected and the object segmented from the background, (c) the light spread image after Gaussian filtering, (d) a 3D plot of the Gaussian filtered light spread, (e) a bi-quadratic polynomial function fitted to the light spread, and (f) the intensity-corrected image generated using the function. Note the intensity differences are very subtle and almost imperceptible so as to not overcompensate, but the differences are appropriate to counter the non-uniform intensity problem.

#### 4. Method

To enable us to compute more accurate surface normal vectors for RTI, the sources of error described in section 3.1 due to non-uniform illumination will be addressed before any estimates of the surface topography are computed. In order to algorithmically correct for non-uniformity in lighting for a given image, the incident illumination in the image scene must first be characterised. It has been discussed in section 3.1 how intensity variations in the light source can be measured photometrically if large enough portions of the image scene are Lambertian. However, this is more difficult since RTI is often performed in arbitrary conditions (often outdoors) when the image scene is very inhomogeneous. Often a given RTI image is comprised of many different objects, and hence surfaces, with different reflectance properties and often at varying distances from the focal plane.

In order to characterise the non-uniform lighting and then compensate for this across the image the intensity profile being emitted from the light source must be iso-

lated as much as possible from the various surfaces in the image. This is described in sections 4.1 and 4.2. Once this light profile has been isolated, a 3D function is fitted to the extracted intensity profile of the light source, and used to marginally brighten the under-lit regions, as described in section 4.3. Finally, the lighting directions are calculated from these fitted intensity profiles; this process is described in section 4.4. Process diagrams for the standard highlight RTI method and the proposed method are shown in Fig. 7 and the algorithm is outlined in algorithm 1.

##### 4.1. Isolating intensity variations due to light source

The first step in this process is to isolate the light spread for each source photograph by subtracting the mean image. This is described by (2), where  $I_c$  is the mean-corrected version of a specific image and is constructed as follows

$$I_c = I - \sum_{n=1}^N \frac{I_n}{N} \quad (2)$$

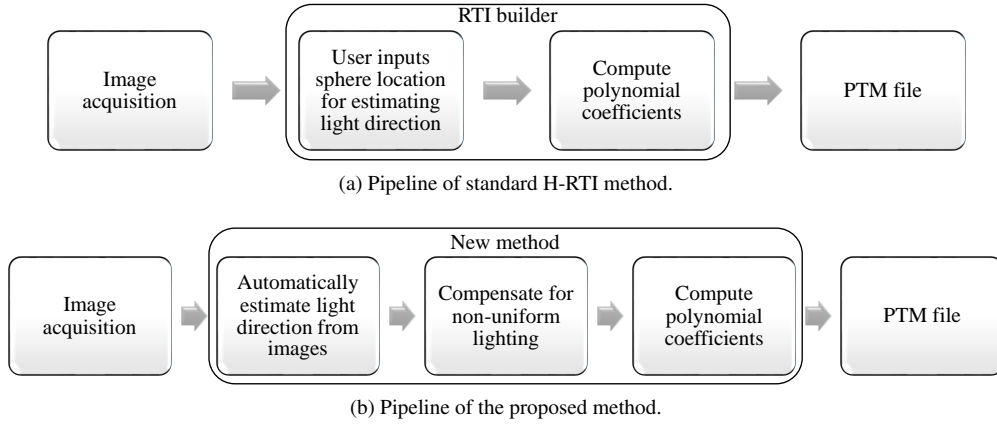


Figure 7: Process diagrams for the standard H-RTI method and the new proposed method.

where  $I$  is a given source image,  $I_n$  is the  $n$ th image in the stack and  $N$  is the total number of images. This leaves us with an image in which the direction of the incoming light is more perceptible, as shown in Fig. 6b. Subtracting the mean image removes most of the salient features of the object from the image so we are left with the incident light. A Gaussian filter is applied to this light intensity to smooth out small irregularities, as shown in Fig. 6c. In Fig. 6 an image analogous to viewing a Lambertian surface has been obtained for our gravestone surface. We can now use this to compensate for non-uniform lighting.

#### 4.2. Optional: segmentation of object from background

In rare circumstances such as the example in Fig. 6, objects are imaged from farther back than would be desired and background objects are present whose reflectance is independent of, or only partially dependent on, the RTI light source (such as bright patches of sky). We provide an optional step here to accommodate for these circumstances, and would like to emphasise that there are a large number of image segmentation methods. Fortunately, most objects are imaged in RTI so as to fill as much of the image frame as possible in order to view the object in higher resolution. The intensity corrections and light direction estimations from the proposed method work without segmentation (see Fig. 12, Fig. 13 and Fig. 15), we simply include this optional step for these special cases.

The automated segmentation process proposed uses Principal Component Analysis (PCA) to separate the foreground and background (Han et al., 2013). The user is presented with the first and second principal components extracted from the image stack, and selects the one that contains the object. The reason for presenting the first two principal components is because often there are multiple objects in the scene when heritage artefacts are imaged in situ and the first and second principal components represent the features that contribute the most to all images. We add the option for the user to manually segment the object by specifying a bounding polygon through a graphical user interface if they have imaged the object from particularly far away.

#### 4.3. Fitting a function to non-uniform light

We choose to fit a bi-quadratic polynomial to each input image, the same function fitted on the reflectance distributions in section 2.4 in the standard highlight RTI method (Malzbender et al., 2001). Note that standard highlight RTI uses the bi-quadratic to estimate the orientation of a given pixel, but here we fit it collectively to all the pixels in a given image for intensity correction. We also opt to fit the bi-quadratic in this instance because it is fast, adequately flexible and does not require interpolation or physical reference spheres to be placed in the scene as per the method proposed by Giachetti et al. (2018). Fitting results for one image of a uniformly flat and diffuse surface are shown in Fig. 8 where the mean-corrected version of



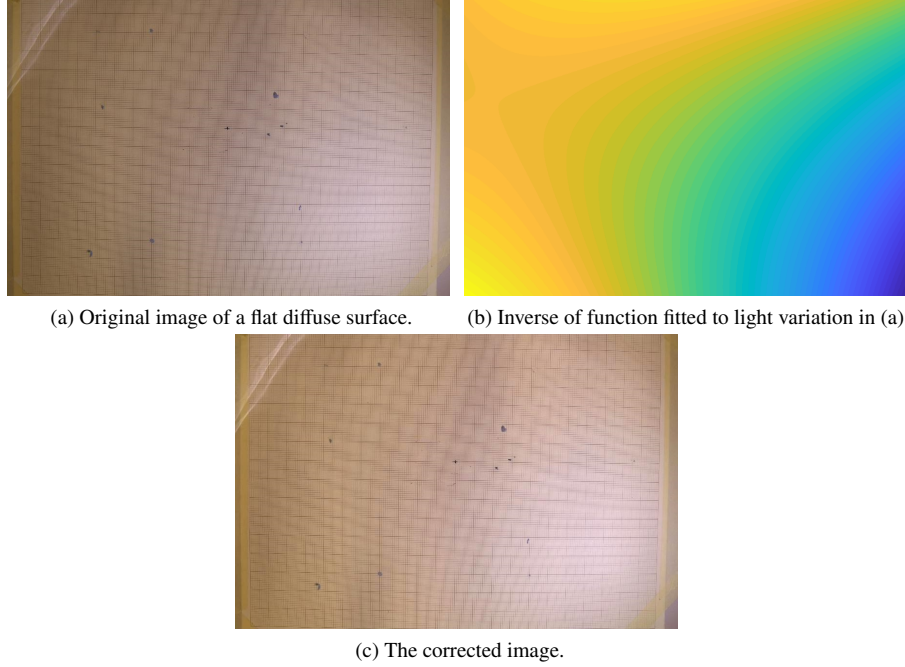


Figure 8: An example of a polynomial function fitted to the isolated light source to correct the light intensity. (a) The original non-uniformly lit image. (b) The inverse of the function fitted to the light variation which is then scaled to correct the image by the process described in section 4.3. (c) The intensity-corrected image generated using the inverse function. Note the darker regions in the original image appear marginally brighter in the corrected image.

the image was generated from the original image to isolate the light source. From here the polynomial function in Fig. 8b was produced which is in turn used to correct the input image in Fig. 8a generating the image shown in Fig. 8c. The full fitting process is shown in Fig. 6, where it can be seen that the intensity variations due to the light source are isolated from the original RGB image. The polynomial is then fitted as shown in Fig. 6e, and this is used to marginally increase the intensity of the underlit region. The differences between the original input image and the corrected image are subtle; non-uniform lighting must not be *over* compensated for. This point is of particular importance since if significantly increase the intensity of a pixel is when it is not receiving light from the light source (due to human error), this would have the effect of amplifying noise. We obtain a marginal brightening factor,  $F$ , by computing the interquartile range,  $\text{IQR}(\cdot)$ , of the

mean-corrected image,  $I_c$  as shown in (3).

$$F = \text{IQR}(I_c) \quad (3)$$

$$P_{\text{corrected}} = (1 + F) P_{\text{original}} \quad (4)$$

This measure of variability ensures that  $F$  is proportional to the variation (or lack thereof) of incoming light in the image. That is to say that less evenly lit images will have a higher interquartile range, and hence have a higher  $F$ . The corrected pixel value,  $P_{\text{corrected}}$ , is given by (4) where  $P_{\text{original}}$  is the original pixel value. This correction is applied to all *RGB* colour channels. We assume that for the majority of images in the image stack all pixels receive light from the light source (albeit some regions more than others). A resulting image which has been corrected for non-uniform intensity is shown in Fig. 6f. It can be seen by inspection that the fitting appears very good at approximating the distribution in light across the image. We will see in section 5 that these assumptions are true for a vari-

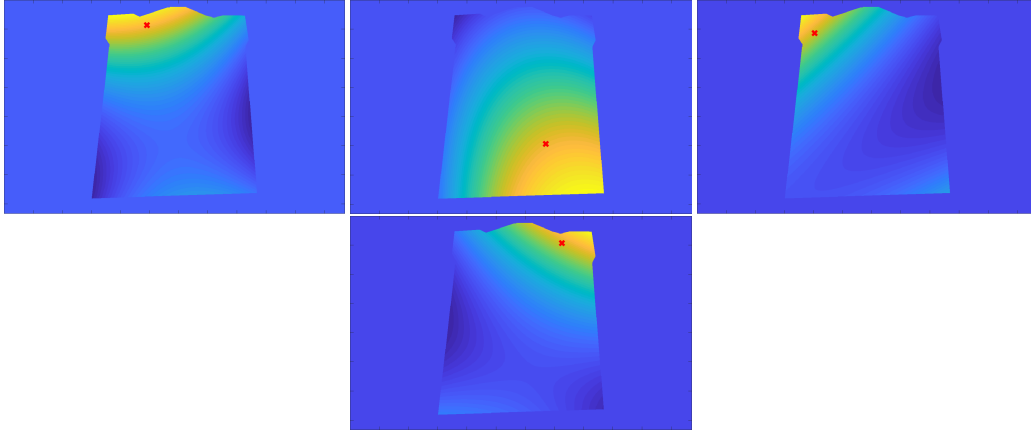


Figure 9: Here we see the automated light direction detection method for the gravestone in Fig. 6. In each image the red cross shows the light direction calculated by the automated method. The function fitted to the variation in intensity is shown in yellow. This function tends to infinity, so we cannot compute its maximum to yield the light direction. If we compute the maximum within the bounds of the image frame we end up at the image periphery. To address this, we compute the centroid of an upper intensity region within the bounds of the image. The limit of this upper intensity is determined by the process outlined in section 4.4.

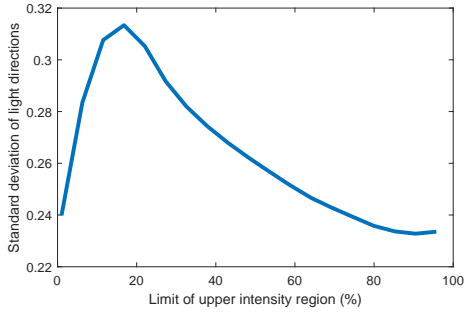


Figure 10: This plot is used in the automatic calculation of light directions as described in detail in section 4.4.

ety of image stacks as these marginal corrections in intensity for RTI purposes prove to be effective in producing more accurate surface normal estimates.

#### 4.4. Automatic light direction estimation

The proposed approach automatically detects the light direction for all images in the image stack, removing the need to insert a reference sphere into the scene. When estimating the light direction in a given image, one may be tempted to compute the coordinates of the maximum point of the fitted function (introduced in section 4.3), extracting the highest value point. This is because one may

be tempted to assume the highest value point of a resulting fit for an image could be approximately equal to where the light source is located, and could yield the direction of the light for a given image. However, it is not possible to compute the global maximum point if the function tends to infinity, as often occurs during fitting. Similarly problematic, if we apply mathematical boundaries to the function (the image borders) and locate the maximum point of the function, we obtain no meaningful information when the function tends towards infinity since the maximum point will always lie somewhere on the border of the image. We propose an efficient method for estimating light direction by computing the centroid (mean coordinates of a group of pixels) of the upper intensity group of points on the fit. This value is much more meaningful, having a more realistic correspondence to the source images than the global maximum. It was found that the optimum percentage of upper intensity for computing this centroid was different for different scenarios, so we produce a robust method which finds this optimum percentage automatically.

Our method samples the centroids computed for various percentage cut-off values and selects the value which yields the highest standard deviation in lighting directions. We do this because a high standard deviation means the estimated light directions are not clustered around the border of the image, and are more evenly distributed, as



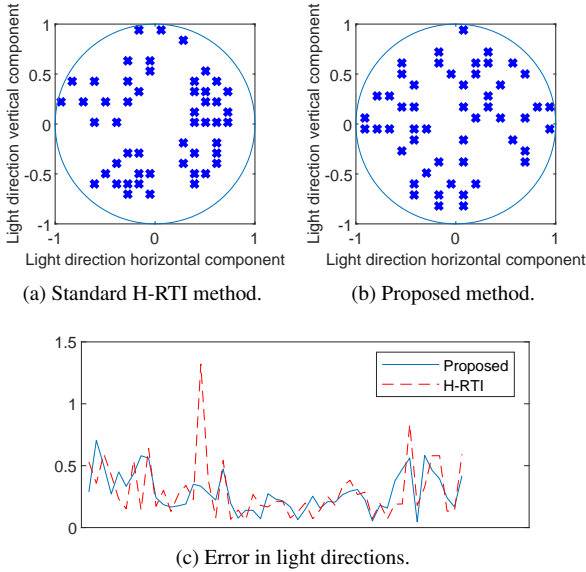


Figure 11: The different light directions computed from the setup shown in Fig. 15a, where the object was imaged in a lighting dome of known lighting directions. Here (a) shows the standard H-RTI method where light directions are more heavily biased to one quadrant due to the bias of the reflective sphere in highlight RTI and (a) shows the proposed method where light directions are more uniformly spread using automatic light direction detection outlined in section 4.4. The errors for the H-RTI and proposed method light directions are shown in (c).

per the true directions of the light source This relationship is shown in Fig. 10, where it can be seen that the standard deviation increases towards 20% upper intensity, then decreases after this optimum percentage.

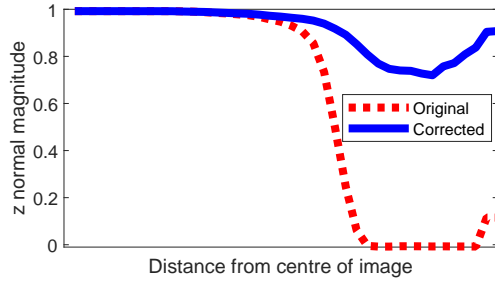
We then take the coordinates of the resulting optimum centroid and normalise with respect to the centre of the image to obtain a lighting direction. This estimation also does not suffer the bias present when using the reference sphere in standard highlight RTI. The result of this process is shown in Fig. 11, which compares the spread of light directions calculated by the standard highlight RTI method with those calculated by our proposed method. It can be seen that the lighting directions generated from the two methods differ greatly, with the reference sphere method suffering from the position bias mentioned in section 3.1 causing the plot of lighting directions to be less evenly distributed, despite the fact that the light source was in fact moved in even coverage around the object. These light directions were computed from the setup

shown in Fig. 15a, where the object was imaged in a lighting dome of known lighting directions. The mean difference between the H-RTI method and the true light direction (shown as the dashed line in Fig. 11c) was found to be 0.2959, and the mean difference between the proposed method and the true light direction (shown as the solid line in Fig. 11c) was found to be 0.2841 meaning the proposed method is more accurate. The sharp peak observed in the H-RTI error is thought to be a result of the sphere being under the shadow of the object, as is often the case for select images in H-RTI, unbeknownst to the user. Using these automatically detected light direction vectors, it is also possible to compute new surface normal data for comparison with the standard highlight RTI method.

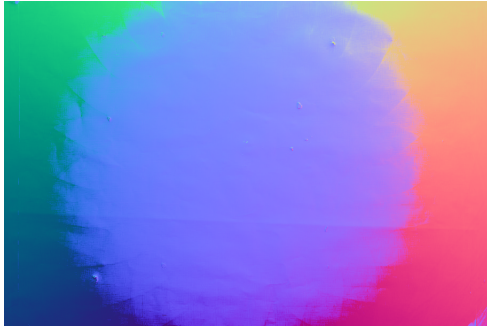
## 5. Experimental Results

We have now fitted a function to the light which allows us to compensate for non-uniform intensity and to extract the light direction automatically. The method generates more accurate incoming light vectors and corrects for non-uniform intensity which results in much more accurate surface normals, as can be seen in Fig. 12 which compares the standard highlight RTI method with the proposed method. Here it can be seen in Fig. 12b that the standard highlight RTI method results in the ring like error effect appearing at the edges due to the periphery of the image receiving disproportionately more light. It can be seen from Fig. 12a that this error has been corrected almost entirely by using the novel method proposed in this paper. Fig. 12a shows this quantitatively, where it can be seen that as we move from the centre of the image out to the right (along the  $x$  axis of this plot) that the  $z$  component of the surface normal drops off from 1 to zero using the standard H-RTI method, but is significantly improved using the proposed method where it only drops down to around 0.8. The true value should be close to 1 across the whole range.

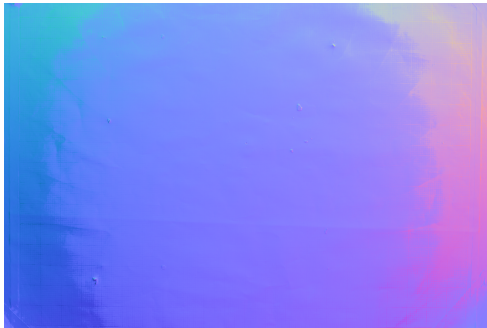
Fig. 13a shows a cylindrical chimney of 2.2m diameter for which surface normals were calculated using the standard highlight RTI method, shown in Fig. 13b, and our proposed method, shown in Fig. 13c. Since the chimney is cylindrical, and constructed from vertical “planks” of concrete, we expect the pixels vertically down the centre of each plank to have the same surface normals, and we can calculate the expected surface normal for each plank.



(a)  $z$  direction surface normal magnitude from centre outwards to the right with data (b) and (c) represented by the red dotted line and the blue solid line respectively.



(b) Standard H-RTI method surface normals of flat surface.

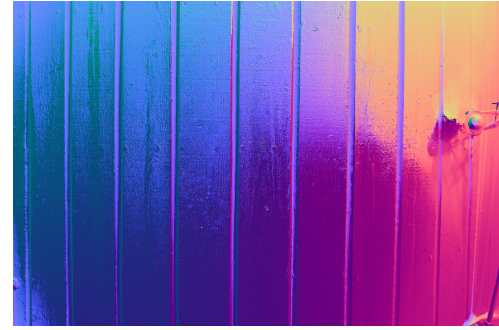


(c) Proposed method surface normals of flat surface.

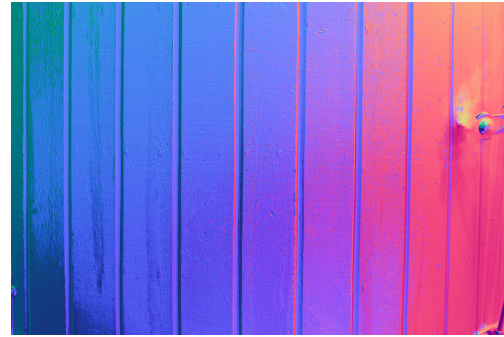
Figure 12: Different surface normal data generated from highlight RTI and the proposed method. (a) The  $z$  surface normal magnitude of the flat surface measured from the centre of image (camera position) outwards to periphery of image. Since the surface is flat and camera-facing the  $z$  normal should measure as a straight line of  $z = 1$  (b) Surface normals generated using traditional RTI. (c) Surface normals generated using non-uniform illumination corrected data and automatic light direction detection (for colour see online version).



(a) Cylindrical chimney of 2.2m diameter.



(b) Standard H-RTI method surface normals.



(c) Proposed method surface normals.

Figure 13: A cylindrical chimney of 2.2m diameter imaged to demonstrate the proposed method on a curved surface. (a) shows the chimney within its background context (this is not one of the RTI images). (b) shows the surface normals of the semi-cylindrical surface calculated using the standard H-RTI method, while (c) shows the same for the proposed method (for colour see online version).

The plots in Fig. 14 summarise the results from this exercise, with the left column of plots showing results for

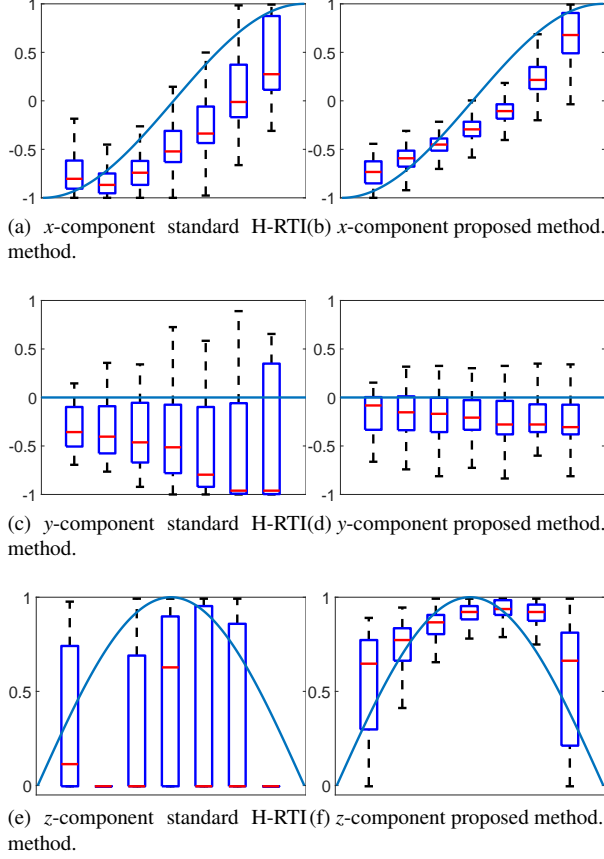


Figure 14: For each of the seven central “planks” of concrete shown in Figs. 13b and 13c, the surface normals for a column of pixels have been used to produce the boxplots shown here. The three rows of plots show the  $x$ ,  $y$  and  $z$  components of the surface normals respectively, where the smooth, continuous line in each plot marks the ground truth. The left column shows results using the standard H-RTI process, while those in the right column are from the proposed method. Each box represents the interquartile range and the centre line inside the box represents the median value; the dashed lines at the top and bottom of each box mark the full range of values.

the  $x$ ,  $y$  and  $z$  components of the surface normals respectively from the standard highlight RTI process, and the right column showing the same for the proposed method. In each plot, the extent of each box marks the 25th to 75th percentiles of the values from the vertical line of pixels from one plank; the line within the box marks the median value. The dashed lines extending vertically from

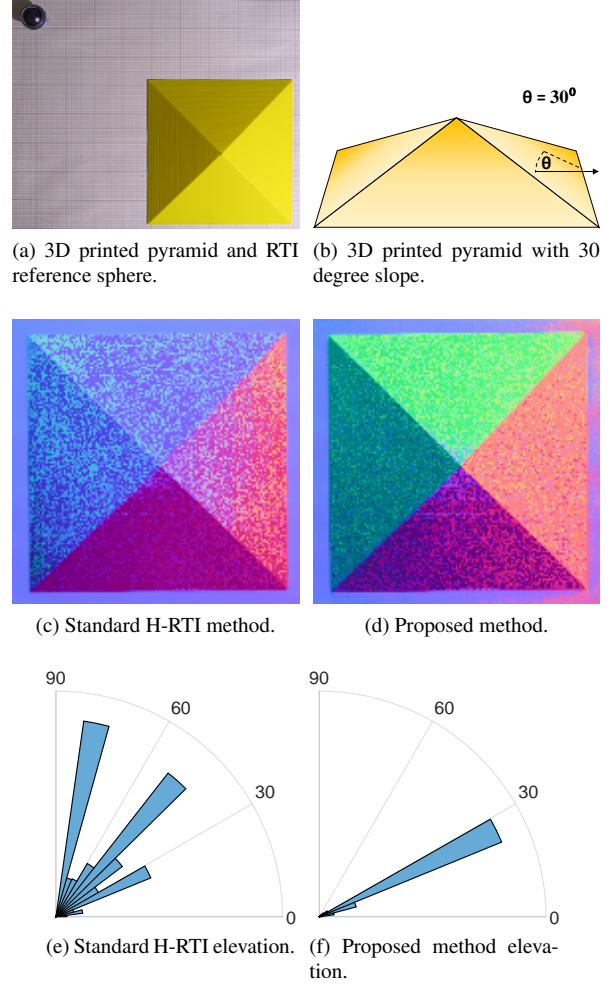


Figure 15: A 3D printed pyramid with a 30 degree slope was produced and imaged to evaluate the proposed method on an object of known geometry. (a) shows the 3D printed pyramid and the RTI reference sphere for producing the standard H-RTI method RTI results (note this image is cropped, the centre of the image is the apex of the pyramid). (c) shows the surface normals of the pyramid faces calculated using the standard H-RTI method, while (d) shows the same for the proposed method (for colour see online version). (e) shows the polar histogram of the elevation angle measured on all four visible faces of the pyramid (ground truth of 30 degrees) for the standard H-RTI method and (f) shows the same for the proposed method.

each box mark the full extent of the values. The single, smooth, solid line in each plot marks the calculated truth. These plots show that the surface normals calculated us-



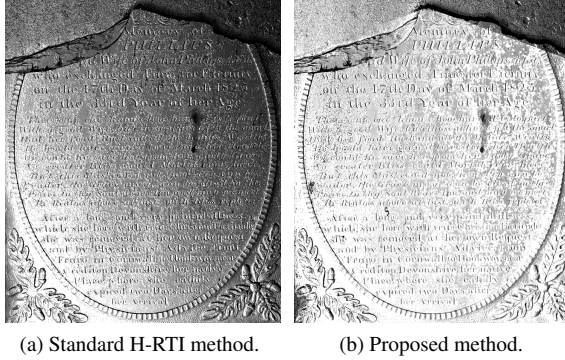


Figure 16: The proposed technique not only results in more accurate surface normal data, but also results in improved readability and uniform lighting in virtual relighting which is visible here for a flat surface which falsely appeared to be curved due to non-uniform lighting. Note these images are lit from the same lighting direction. (a) Specular enhancement for uncorrected RTI image stack showing non-uniform lighting. (b) Virtual relighting specular enhancement for corrected RTI image stack showing uniform lighting using the proposed method. This also increases readability.

ing the proposed method are significantly more accurate than those from the standard H-RTI method, with smaller spreads of values that are closer to the ground truth. This is especially obvious for the  $y$  and  $z$  components.

In order to evaluate the light directions estimated and the surface normals generated by the proposed method we image an object of known geometry under known lighting conditions. A 3D printed pyramid with a 30 degree slope was produced and imaged. Fig. 15a shows the 3D printed pyramid and the RTI reference sphere used to produce the standard highlight RTI results for comparison. Fig. 15c shows the surface normals of the pyramid faces calculated using the standard highlight RTI method, while Fig. 15d shows the same for the proposed method (for colour see online version). Polar histograms of the elevation angle measured on all four visible faces of the pyramid (ground truth of 30 degrees) for the standard H-RTI method and the proposed method are shown respectively in Fig. 15e and 15f. It is clear from inspection that the proposed RTI method produces results much closer to the true elevation angle (of 30 degrees) of the pyramid faces, yielding a mean of 25.19 degrees with a standard deviation of 6.20 degrees, whereas the standard highlight RTI method yields a mean of 51.1 degrees with a standard deviation of 20.21 degrees.

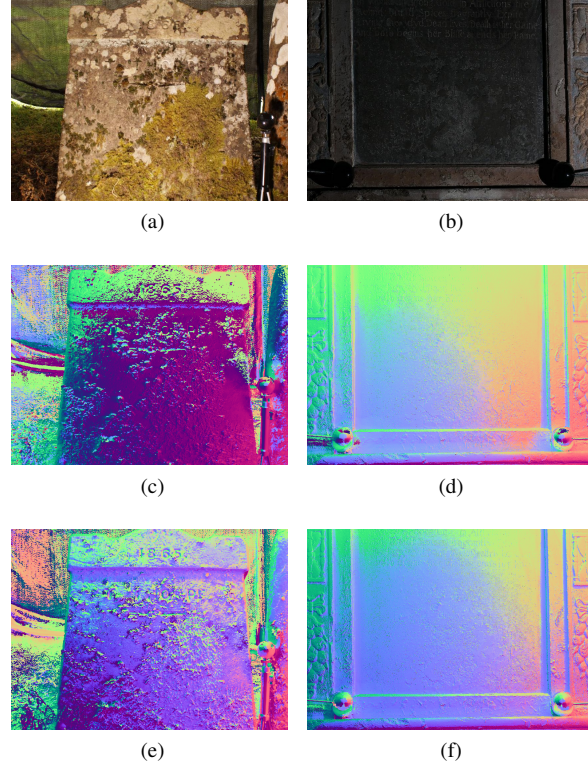


Figure 17: Here we have retrofitted the proposed method to existing RTI image stacks. (a) shows a colour image of the gravestone featuring in Fig. 6. In (c) we see the surface normals measured for the same gravestone using the standard H-RTI method and (e) shows the surface normals measured using the proposed method. (b) shows a colour image of a flat wall memorial. (d) shows the surface normals calculated for the same wall memorial using the standard H-RTI method and (f) shows the surface normals calculated using the proposed method.(for colour see online version).

The proposed method was also retrofitted to noisy and imperfect pre-existing RTI image stacks in order to test its robustness. The results of this retrofitting are shown in Fig. 17. The retrofitting of the proposed method to the gravestone shown previously in Fig. 6 is shown in Fig. 17e and the standard H-RTI method results are also shown in Fig. 17c. Here it is shown that the traditional RTI method with the reference sphere generated such poor surface normal data of this generally flat gravestone facing the camera ( $z$  direction) that large portions of this surface appear green ( $y$  direction i.e. upwards). Despite the grave surface in Fig. 17a consisting of different textures (stone

and various mosses) the surface normals image should appear almost uniform in colour (blue/purple) because the gravestone is flat and facing directly at the camera (and so are the various mosses by and large). These inaccuracies apparent in surface normal data caused by non-uniform illumination are visible from the variation in colour of the surface normal image in Fig. 17c. Retrofitting was also carried out on a pre-existing image stack of a (also mostly flat) wall memorial shown in 17b, 17d and 17f. It is shown in Fig. 17f that the proposed method surface normals appear more uniform and camera-facing for the same RTI image stack using the lighting uniformity corrections and automated light direction detection. It is clear here that the surface normals represent a more realistic, flat and uniform surface. These image stacks are typical of highlight RTI performed on medium to large objects, where for the most-part the lighting has been uniform but occasionally there over-lit regions and completely under-lit images. Despite this, it was found that the method is robust and improves the accuracy of the surface normal data.

Fig. 16 shows how the proposed corrections from section 4 additionally benefit the virtual relighting obtained from RTI. Fig. 16a shows a flat surface under RTI virtual relighting in which the standard H-RTI method has yielded a surface which appears, erroneously, to bulge due to non-uniform lighting. Fig. 16b shows the same object, relit from the same virtual lighting direction, but processed using the proposed method. The surface appears much flatter, and is, therefore, much more uniformly lit, which increases the legibility of the inscriptions.

## 6. Conclusions

The results discussed in section 5 show marked improvements in surface normal estimation for highlight RTI using the proposed new method which, as is shown in Fig. 17, can also be retrofitted to existing highlight RTI image stacks. RTI for larger objects suffers the most from non-uniformity in lighting, since for small items, such as coins, the lighting can be considered approximately parallel as the deviation in incident light angle across the image plane is much smaller. The proposed method more accurately estimates the light directions (see section 4.4) and does not require the addition of reference spheres or 3D printed calibration devices into the scene (note the 3D

printed pyramid is for evaluation purposes). We keep the RTI process simple and inexpensive for non imaging science/photography professionals whilst markedly improving the results for surface normal generation compared with highlight RTI when performed on larger objects and objects of known geometry such as in Fig. 15.

The significant improvement in uniform lighting evident in Fig. 16 during virtual relighting shows that the method is also relevant to conservationists who are more interested in readability of inscriptions than quantitative RTI.

These more quantitatively accurate surface normals mean that RTI could become more ubiquitous in cases where laser scanning would usually be used. This technique is not invulnerable to very messy RTI image stacks, where the light source has largely missed the centre of the object for a significant proportion of the image stack, but is more robust than traditional RTI.

## References

- Agrawal, A., Raskar, R., and Chellappa, R. (2006). What is the range of surface reconstructions from a gradient field? In *European Conference on Computer Vision*, pages 578–591. Springer.
- Boute, R., Hupkes, M., Kollaard, N., Wouda, S., Seymour, K., and ten Wolde, L. (2018). Revisiting reflectance transformation imaging (RTI): A tool for monitoring and evaluating conservation treatments. *IOP Conference Series: Materials Science and Engineering*, 364:012060.
- Cultural Heritage Imaging (2016). *Reflectance Transformation Imaging (RTI)*, (accessed September 6, 2018).
- Edwin, F. (1977). *Geometrical considerations and nomenclature for reflectance*, volume 160. US Department of Commerce, National Bureau of Standards.
- Gautron, P., Krivanek, J., Pattanaik, S. N., and Bouatouch, K. (2004). A novel hemispherical basis for accurate and efficient rendering. *Rendering Techniques*, 2004:321–330.
- Giachetti, A., Ciortan, I. M., Daffara, C., Marchioro, G., Pintus, R., and Gobbetti, E. (2018). A novel framework

for highlight reflectance transformation imaging. *Computer Vision and Image Understanding*, 168:118–131.

Hammer, Ø., Bengtson, S., Malzbender, T., and Gelb, D. (2002). Imaging fossils using reflectance transformation and interactive manipulation of virtual light sources. *Palaeontologia Electronica*, 5(1):1–9.

Han, Y., Feng, X.-C., and Baci, G. (2013). Variational and PCA based natural image segmentation. *Pattern Recognition*, 46(7):1971–1984.

Hewlett-Packard (2001). *PTM fitter*; (accessed September 6, 2018).

Huang, X., Walton, M., Bearman, G., and Cossairt, O. (2015). Near light correction for image relighting and 3D shape recovery. In *Digital Heritage, 2015*, volume 1, pages 215–222. IEEE.

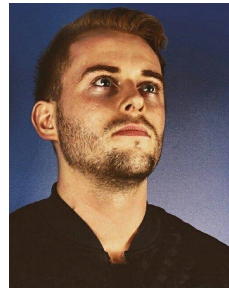
Malzbender, T., Gelb, D., and Wolters, H. (2001). Polynomial texture maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 519–528. ACM.

Mudge, M., Malzbender, T., Schroer, C., and Lum, M. (2006). New reflection transformation imaging methods for rock art and multiple-viewpoint display. In *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, volume 6, pages 195–202.

Pitard, G., Le Goïc, G., Favrelière, H., Samper, S., Desage, S.-F., and Pillet, M. (2015). Discrete modal decomposition for surface appearance modelling and rendering. In *Optical Measurement Systems for Industrial Inspection IX*, volume 9525, pages 952523:1–10. International Society for Optics and Photonics.

Winnemöller, H., Mohan, A., Tumblin, J., and Gooch, B. (2005). Light waving: Estimating light positions from photographs alone. *Computer Graphics Forum*, 24(3):433–438.

## 7. Vitae



**Matthew McGuigan** is a PhD student in Computer Science at the University of Exeter. He received his BSc in Physics and Astronomy from the University of Glasgow (UoG) and his MSc in Sensor and Imaging Systems jointly from the UoG and the University of Edinburgh. His research interests include machine learning, computer vision, optics and astronomy.



**Jacqueline Christmas** is a Senior Lecturer in Computer Science at the University of Exeter. Her research interests include machine learning for intelligent image and video understanding, Bayesian modelling and variational approximation. She heads the RTI Group at the University of Exeter.